

Shared Investor Networks: Linking Microsoft (*MSFT*) to the Market

Anaya Gandhi

October 2024

Introduction

This report examines investor behaviour and their relationship with major stocks in the market with a focus on the Microsoft (MSFT) stock. The analysis spreads over a range of tasks such as creating bipartite graphs and ego networks for particular stocks, helping in understanding specific patterns and relationships to examine investor traits and behaviours. The dataset looks at data over the years 2016 to 2023, which includes the period of the pandemic. The aim is to identify trends and possible influences on the Microsoft stock to make informed and strategic investment decisions and to investigate how investor behaviour forms connections between MSFT and other significant stocks through the lens of shared investors. While this analysis focuses on a broader temporal dataset encompassing various stocks across different sectors, MSFT is the focal point. The project tracks how changes in investor portfolios during critical periods, such as the pandemic, impact MSFT's performance and its connections with other stocks. Through this analysis, DBBA Capital aims to identify trends and potential influences on MSFT that could inform more strategic and informed investment decisions in the future. In this report, while the nodes denote investors and their investments, the edges represent the relationship between these nodes. The degree of the nodes depict the number of stocks that share at least one common investor. That is, it shows how many stocks are linked together based on common investors, highlighting the common interests and patterns among these investors. The degree of the nodes are represented by a gradient scale of the colour blue. As we go from light to darker shades of blue, the degree of the nodes increases. On the other hand, the weight of these edges indicate the number of periods (Quarters) an investor has invested in a particular stock between 2016-2023, determining the strength of the relationship over the years. The edge weights are represented by a gradient scale from yellow to red where yellow represents the lowest edge weight. Increases in the value as the scale darkens, emphasises the degree of association over various graphs.

Task 1 - Constructing a Bipartite Network

In a Bipartite network, the nodes are separated into two sets and only linked to nodes in the opposite set. In the graphs below we have two sets (Investors and investment) where the investors are linked with their investment stocks. A single data frame was created which extracted data from all multiple.xlsx files covering data from the 2016-2023 period. The columns con-

tained information regarding the investor, the year, the quarter and the stock in which they invested. This data, stored in a long format, was used to create the bipartite network. The weight of the edges in the graph determine the number of times an investor invested in that stock, helping track investor behaviour and patterns over time.

Task 1.1 - Visualising the Network

Using the dataset provided, a bipartite graph was created to understand the relationship between investors and their investments. The graph in Figure 1 was made by iterating through each row in the dataset, and it has two sets of nodes-investors and nodes, which are linked together with an edge. There are a total of 1606 nodes in the graph out of which 77 investor nodes are represented by the colour pink and 1529 stock nodes shown by the colour wheat. The edge weight increases each time an investor invests in the stock. In the graph below, the edge weight is further normalised between 0 and 1 and depicted by a gradient, which helps differentiate between investors who make large investments and those who do not. The plot highlights the complex nature of the connections between the investors and their investments, implying that several investors have a diverse portfolio and have invested in multiple companies.

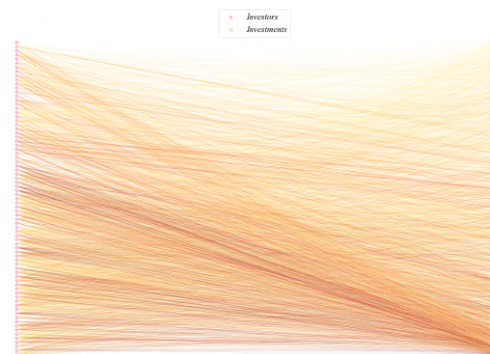


Figure 1: *Weighted Bipartite Ego Network of Investors and Investments from 2016 to 2023.*

A histogram of normalised weight is created to analyze the strength and connectivity of the above plot. The histogram in Figure 2 reveals that majority of the investors have a small stock portfolio (indicated by low edge weights), while a smaller group of investors have a big portfolio and invest in certain stocks consistently over time periods. This graph highlights the difference in investment strategies among the various investors.

The difference in investors' approach towards stock investments, to see if they are inclined towards high turnover, short term investments or prefer stable, long-term investments in particular stocks can be observed by looking at both the above graphs.

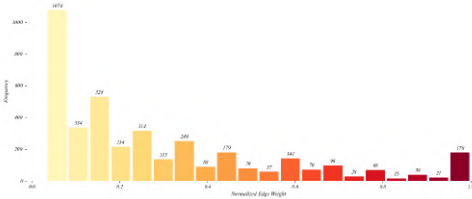


Figure 2: Histogram of Normalized Edge Weights

Task 1.2 - Bipartite Ego Network

In Task 1.2, Microsoft's (MSFT) investor landscape is examined by creating an ego network centred around the company. In this network, MSFT is connected to all the other investors who have held its stock, which emphasises the extent of their investments. Like a central node, MSFT is linked to all its investors through edges that show their connection. The weight of each edge depicts the number of time periods the investor held MSFT shares, emphasising each investor's consistency and commitment. In this visualisation, MSFT is depicted as a prominent red node, while the connected investors are shown as peach-coloured nodes. The edges are shown with the help of grey lines and the labels depict the frequency of investments between MSFT and each investor. This graph helps in understanding the direct relationships between MSFT and its investors while identifying those connections that have remained committed to the company over a longer period.

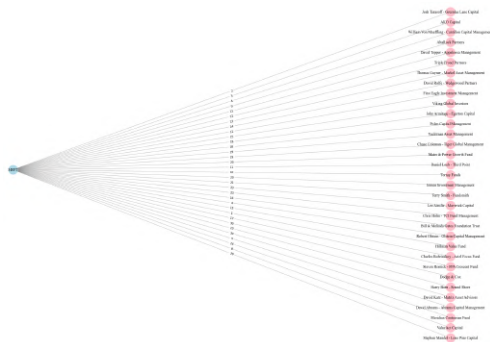


Figure 3: Weighted Bipartite Ego Network of MSFT from 2016 to 2023.

Table 1 depicts a diverse range of investors and asset managers such as, hedge funds, mutual funds, charitable trusts, and activist investment firms. Many investors, like Dodge and Cox, Hillman Value Fund, Yacktman Asset Management, and Robert Olstein, follow a value investing strategy and target undervalued stocks with strong cash flows. For example, investors such as the Bill and Melinda Gates Foundation Trust and Jensen Investment Management invest in established large-cap companies,

thereby prioritising capital preservation and stability over aggressive growth (Gates Foundation Portfolio Report, 2023). A growth investing strategy is used by firms like Maverick Capital, Fundsmith, Polen Capital, and Wedgewood Partners to capitalise on future growth potential. Hedge funds like Tiger Global Management often invest in the broader technology ecosystems (Goldman Sachs Technology Sector Report, 2023) while activist investors such as Third Point and ValueAct Capital aim to unlock shareholder value through strategic influence (Bloomberg, 2024). Value investing appears to be the prevailing strategy used by those investing in Microsoft (MSFT). Growth investing is also dominant in sectors like healthcare and technology in accordance with MSFT's industry. In addition, some of the most important players use activist strategies to try and create more value, while the other investors prefer a passive long term approach. This proves that investors see the MSFT stock for its stability as well as its long term growth potential as it attracted a wide range of investors

Task 2 - Network Projections

Task 2.1 - Creating a Stock-Side Network Projection

A network projection refers to a transformation process used in bipartite networks (or two-mode networks) to create a simpler, one-mode network. In a simple bipartite network, there are two types of nodes, and connections only exist between nodes of different types. A network projection reduces this bipartite structure to focus on one type of node, connecting them based on shared associations with the second type of node. In this task we created a network projection on the stock side of the bipartite network, the nodes in this network represent stocks and the edges act as a link between two stocks. This link portrays that two stocks share at least one common investor and the colour of the edge determines the degree of connection between the two nodes. The degree denotes how many different stock shares have overlapping investors to help identify particular investor behaviours. In the plots the colour of the edges represents the degree, i.e darker the colour higher is the number of shared investors. The stock-side network projection was portrayed in two different layouts - Circular and Spring. In Figure 4 a circular layout is shown which helps understand the overall structure of the network. Figure 5 emphasises a spring layout which helps identify clusters and the density of connections. The DBBA capital could use this to identify certain stocks that are perceived as co-investments by investors. For example, a strong link between a stock could show market trust in both companies, suggesting a possible strategic opportunity for future investment. The stocks with the highest number of common investors that are frequently co-invested with other stocks are MSFT (231), GOOG (228), GOOGL (222), AMZN (209), and META (206). The stocks with the least number of common investors are HGV (4), WRLD (4), SWBI (3), WGO (3), and AMR (1), showing a lower market interest.

Investor	Number of Investments	Investment Firm Type
David Katz - Matrix Asset Advisors	30	Asset Management
Dodge & Cox	30	Mutual Fund
Stephen Mandel - Lone Pine Capital	29	Hedge Fund
Chris Hohn - TCI Fund Management	23	Hedge Fund
Lee Ainslie - Maverick Capital	23	Hedge Fund
Terry Smith - Fundsmith	21	Asset Management
Mairs & Power Growth Fund	20	Mutual Fund
Jensen Investment Management	20	Mutual Fund
Yacktman Asset Management	19	Asset Management
Chase Coleman - Tiger Global Management	19	Hedge Fund
Polen Capital Management	18	Asset Management
John Armitage - Egerton Capital	18	Hedge Fund
Steven Romick - FPA Crescent Fund	17	Mutual Fund
Bill & Melinda Gates Foundation Trust	16	Charitable Trust
Viking Global Investors	15	Hedge Fund
Hillman Value Fund	15	Mutual Fund
David Rolfe - Wedgewood Partners	14	Asset Management
Meridian Contrarian Fund	14	Asset Management
Thomas Gayner - Markel Asset Management	13	Asset Management
Triple Frond Partners	12	Hedge Fund
First Eagle Investment Management	11	Asset Management
Daniel Loeb - Third Point	11	Hedge Fund
David Tepper - Appaloosa Management	11	Hedge Fund
Torrey Funds	10	Asset Management
Harry Burn - Sound Shore	10	Asset Management
AltaRock Partners	9	Hedge Fund
ValueAct Capital	8	Activist Investment
David Abrams - Abrams Capital Management	5	Hedge Fund
Robert Olstein - Olstein Capital Management	4	Asset Management
William Von Mueffling - Cantillon Capital Management	4	Hedge Fund
AKO Capital	3	Hedge Fund
Josh Tarasoff - Greenlea Lane Capital	2	Hedge Fund
Charles Bobrinskoy - Ariel Focus Fund	1	Mutual Fund

Table 1: *Investors and Number of Investments*



Figure 4: Stock-Side Network Projection: Circular layout

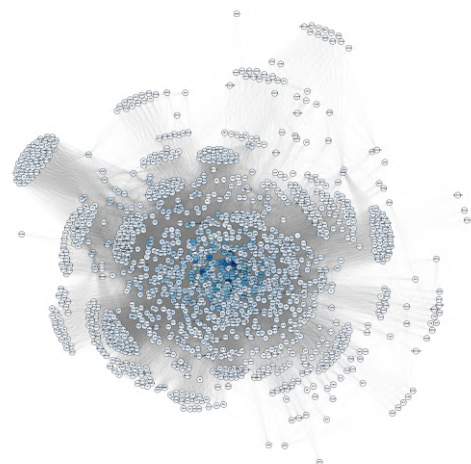


Figure 5: *Stock-Side Network Projection: Spring layout*

Task 2.2 - Network Comparison

Bipartite networks represent both, stocks and investors to emphasise the relationship between them and to understand investor activity across different stocks and analyze their portfolios. The network projection plots only nodes and simplifies the network to highlight the relationship between companies having common investors. The differences between the two projections are summarized in Table 2. Both network types are important and required to help the DBBA Capital Identify important investors and their stock interests. This would help with specific strategies to attract those investors. These networks can help investors make informed decisions and diversify their portfolios across related sectors . The bipartite network helps identify major investors and their strategies which is very useful to understand their behaviour and choices. The network projection focuses on clusters of stocks invested together, which helps to understand market trends

Bipartite Network	Projection Network
Contains nodes representing both investors and stocks.	Contains nodes representing stocks only.
Represents direct relationships between investors and stocks.	Represents connections between stocks based on shared investors.
Provides insights into key investors and their individual strategies.	Highlights relationships between stocks through co-investment patterns.
Useful for understanding individual investment behaviors.	Useful for identifying clusters of stocks and market trends.
Helps in identifying influential investors for targeted strategies at DBBA Capital.	Guides diversification and cross-investment opportunities at DBBA Capital.

Table 2: Comparison between Bipartite Network and Projection Network

Task 2.3 - Ego Network Comparison

The unipartite ego network for Microsoft (MSFT) was created using the stock-side network projection from Task 2.1. This network portrays all the connections between MSFT and other stocks that have common investors. The MSFT stock is the central node highlighted with a dark blue colour as it has the highest degree. The network was visualised using two circular as well as spring layouts. The former portrayed an overview of the entire network while the latter highlighted the relationship strength and clusters. The difference between the two graphs is very evident. In the bipartite network, MSFT is linked to the variety of investors that have invested in the stock. In the unipartite network, MSFT is connected to other stocks to show the relationship between MSFT and other companies based on common investors. The neighbours in the unipartite network are

greater than the bipartite network. This is because the unipartite network focuses more on the relationship between stocks sharing common investors which creates more links as compared to individual investor relationships. Since the network has neighbours which represent stocks that are commonly co-invested with MSFT, a pattern in investor preferences can be gauged. The network shows that MSFT has strong connections with other technology related stocks such as AMZ (158), GOOGL (150), GOOG (142) and META (135). On the other hand a stock in the healthcare sector TRUP (11) and in the Financial sector BUR (9) have very few common investors. This implies that investors often include MSFT as a part of the broader technology portfolio. This also shows that investors believe that MSFT and related stocks have similar risks and rewards, and hence follow consistent trends. With this graph and information, DBBA Capital can identify sectors closely linked with MSFT and make educated decisions to predict trends and patterns in the market.



Figure 6: Unipartite ego network for Microsoft: Circular Layout

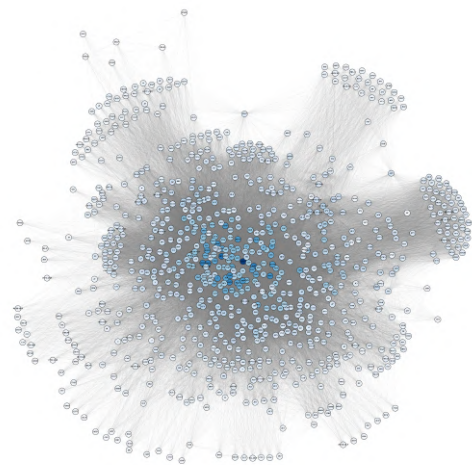


Figure 7: Unipartite ego network for Microsoft: Spring Layout

Statistic	Mean	Maximum	Minimum	Standard Deviation
Number of Nodes	586.90	648.00	496.00	60.02
Number of Links	8720.10	9823.00	6948.00	1185.91
Density	0.0509	0.0575	0.0462	0.0037
Average Clustering Coefficient	0.8225	0.8405	0.8078	0.0077
Average Degree	29.61	31.64	27.42	1.17
Average Strength	32.74	36.05	29.12	2.13
Assortativity	-0.0272	0.0232	-0.0677	0.0309

Table 3: *Whole Network Summary Statistics*

Statistic	Mean	Maximum	Minimum	Standard Deviation
Number of Nodes	187.97	240.00	104.00	43.30
Number of Links	2998.93	4100.00	1305.00	960.82
Density	0.1723	0.2437	0.1356	0.0266
Average Clustering Coefficient	0.7720	0.8257	0.7339	0.0228
Average Degree	31.17	36.13	25.10	3.69
Average Strength	38.53	47.14	28.56	6.25
Assortativity	-0.1185	-0.0475	-0.1770	0.0461

Table 4: *MSFT Ego Network Summary Statistics*

1 Task 3 - Basic Network Analysis

1.1 Task 3.1 - Network Statistics

In Figure 8 a network projection on the stock side for each quarter was plotted to note the change in investment patterns over time. This helps observe how stocks are linked with common investors over time and identify a shift in trends. The MSFT ego network for each quarter is highlighted by nodes having the colour beige. There is a consistent dense cluster around MSFT over time which portrays strong connections and confident investments among the investors and stocks, the slight variations in these clusters highlight changes in co-investment preferences. This plot also highlights how MSFT fits into the broader market. Table 4 consists of summary statistics (i.e mean, maximum, minimum and standard deviation) of various computed quantities for the whole network as well as the MSFT ego network. The quantities computed were -

1. Number of nodes
2. Number of links
3. Density
4. Average clustering coefficient
5. Average degrees
6. Average strength
7. Assortativity

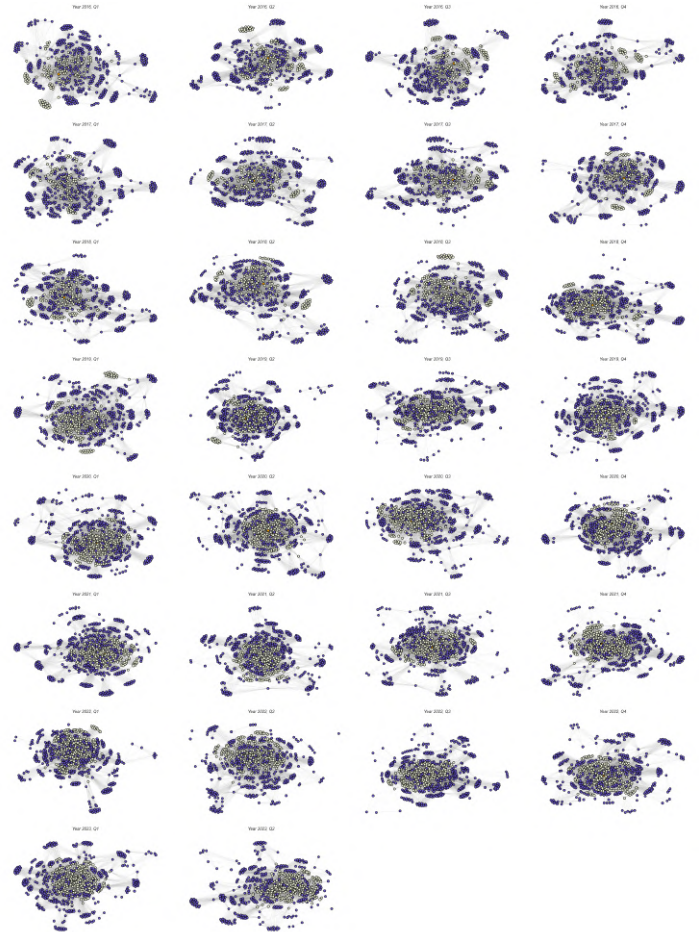


Figure 8: *Weighted Bipartite Ego Network of MSFT from 2016 to 2023.*

Task 3.2 - Discussion

The whole network has an average of 587 nodes and 8720 links while MSFT has only 188 nodes and 2999 links. This difference is expected because MSFT network focuses only on stocks that share a common investor while the whole network represents all stocks. The density of the MSFT ego-network is higher, having an average of 0.1723 compared to 0.0509 for the whole network. This indicates that stocks related to MSFT are more interconnected and investors who invest in MSFT tend to co-invest in stocks within the same group, hence making it denser. Such high-density clusters imply a high level of investor confidence in MSFT related stocks. The average clustering coefficient is lower in the MSFT ego network compared to the whole network. The clustering coefficient measures the likelihood that a stock's neighbours are also connected to each other. A lower clustering coefficient in the MSFT ego-network indicates that although the density is high these links do not always form clusters. This suggests that while MSFT serves as a central point, the other stocks in its ego network are not strongly related to each other. The average degree is slightly higher for the MSFT ego network (31.7) than for the whole network (29.61). Similarly, the average strength or the MSFT ego network (38.53) than for the whole network (32.74), implying that, on average, the stocks connected to MSFT have more connections and stronger relationships compared to the broader market. This suggests that MSFT is part of an important group of stocks for investors, often tied to other well-known investment options; a pattern that suggests investor confidence and a consistent focus on technology and related sectors. The average clustering coefficient for the whole network is 0.8225 and it is slightly lower for the MSFT ego network at 0.7720. The lower value in the MSFT ego network suggest that although there are multiple common investors between the stocks, they do not form a tight cluster. This /indicates a range of diverse investments, which could reveal synergies or complementarities that aren't immediately visible. The whole network as well the the MSFT network have negative assortativity values. The whole network averages -0.0272 and the MSFT ego network shows a more negative value of -0.1185. This implies a disassortative pattern in both the networks where stocks with high number of connections are often connected with stocks having fewer connections. This negative assortativity which is more evident in the MSFT network, which suggests that investors holding MSFT invest in both popular and less popular stocks to have a balanced portfolio and lower their risk to reward ratio. From a strategic approach, understanding the density of MSFT's ego-network provides insights into investor behaviour. The higher density suggests that DBBA Capital should consider co-investment opportunities in stocks closely associated with MSFT, as investors show a pattern of holding these stocks together. The lower clustering coefficient means that while MSFT is linked to many other stocks, which aren't necessarily forming tight knit clusters. This could be a good opportunity for DBBA Capital to look deeper into potential synergies that aren't immediately obvious. The higher average degree and strength show that stocks associated with MSFT tend to have strong co-investment ties, making them potentially

solid choices to invest in alongside MSFT. Lastly, the negative assortativity in both networks points to investors using a diversified strategy—mixing well-known, popular stocks with lesser-known ones. DBBA Capital could take a similar approach to manage risk while gaining exposure to both big names and emerging opportunities.

Task 4 - Changes in Network Statistics During the Pandemic

Task 4.1 - Temporal Evolution of Statistics

The temporal analysis sheds light on metrics like the Number of Nodes, Density, Average Degree, and Assortativity, which play an important role to understand invest behaviour during challenging periods like the pandemic. Quantities like Number of Nodes and Density offer a view of how the network either grew or fell, emphasizing the confidence investors had in investing in new stocks. Average Degree and Average Strength reveal how investors focus on a small number of stocks, especially Bluechip stocks, during uncertain times. Assortativity indicates the diversity in investment strategies and shows whether investors were leaning towards high-profile stocks or spreading their bets across other less popular stocks. During the pandemic, we observed an increase in density and average strength within the MSFT ego-network, this tell us that MSFT and its related stocks became a go-to choice for many investors seeking stability and security. The number of nodes declines which shows how investors focused on trusted and well-known stocks. The presence of negative assortativity during this time shows a more balanced approach to risk, with investors combining well-known stocks with less popular ones to minimize risk. These findings point to how much the pandemic influenced investor decisions regarding MSFT. For DBBA Capital, recognizing these patterns can be really helpful in planning resources and strategies in unpredictable times.

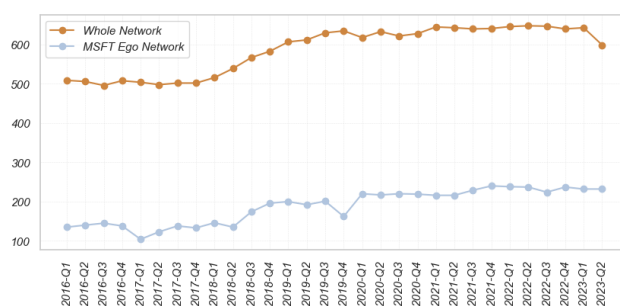


Figure 9: Number of Nodes in the Bipartite Ego Network of MSFT from 2016 to 2023.

This graph represents the evolution of the number of nodes in the bipartite ego network of MSFT, showcasing the investors connected to MSFT over time. The number of nodes represents unique stocks in each network. In the period leading up to the pandemic, there was a steady increase in the number of nodes in

both the networks which indicates the investors were diversifying their portfolios. There is a decrease in the number of nodes in the MSFT network during the start of the pandemic, due to future uncertainty. Post the pandemic i.e After March 2021 there is a slight increase in the number of nodes in the whole networks. Subsequently the number of nodes in the MSFT network plateaued, showing that many investors continued to invest in MSFT related stocks

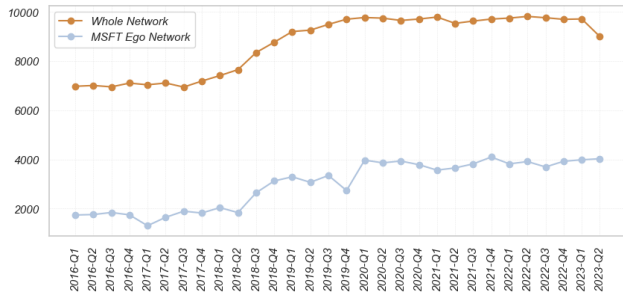


Figure 10: Number of Links in the Bipartite Ego Network of MSFT from 2016 to 2023.

The figure illustrates the number of links in the MSFT ego network, indicating how the number of investments by the connected investors changed over time. The number of links represents the number of shared investors between stocks. The number of links in both the networks show an increasing trend and hence proves that investors are broadening their investments which is in turn leading to this increase. During the pandemic there was a dip in the MSFT ego network because the investors became more conservative and maintained their portfolio rather than investing more. In this same period the whole network saw a steady increase which shows that even though some investors pulled back others saw this as an opportunity to keep investing. However after the March 2021 the number of links once again increased to show investors confidence in the market.

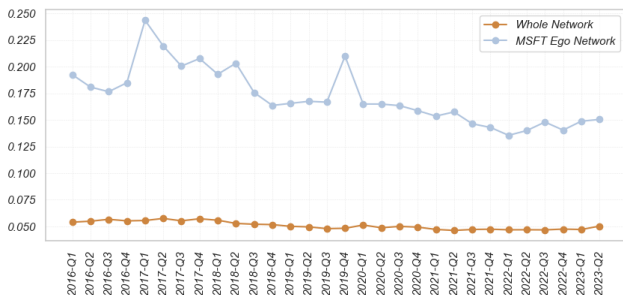


Figure 11: Density of the Bipartite Ego Network of MSFT from 2016 to 2023.

The density of the network reflects how well connected the network is relative to its size, providing insights into the interconnectedness of MSFT’s investors. In the whole network the density was low which portrays that there is a vast number of stocks, but very few connections between them. The MSFT network has higher density which proves that the stocks are more

closely related through common investors. Post the pandemic there is a slight decrease as new stocks which were added to the market once again made the network less dense but more diverse.

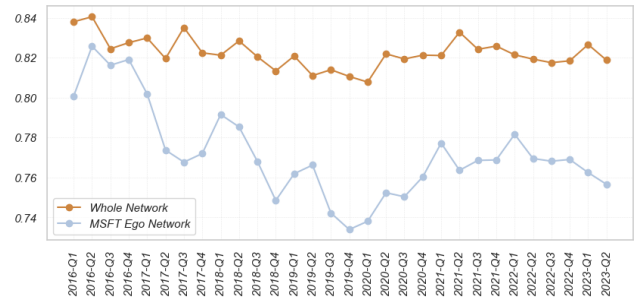


Figure 12: Average Clustering Coefficient in the Bipartite Ego Network of MSFT from 2016 to 2023.

This figure shows the average clustering coefficient, which measures the likelihood that two investors connected to MSFT are also connected to each other. The whole network showed a higher clustering coefficient, which portrays that there were more localized clusters of stocks sharing investors as compared to the MSFT ego network. During the pandemic, the clustering coefficient of the MSFT network dropped which shows that even though MSFT is linked to several stocks it didn’t form tight clusters. The clustering coefficient increased slightly post March 2021 showing that investors were investing in more stocks increasing the connection as the market recovered.

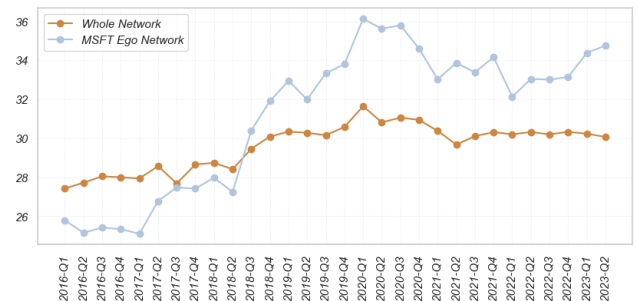


Figure 13: Average Degree in the Bipartite Ego Network of MSFT from 2016 to 2023.

The average degree represents the average number of connections each investor in the network has, indicating the degree of activity among MSFT investors. The average degree in both the networks were relatively stable at the beginning with a slight increases reflecting a balanced approach. During the pandemic, the average degree of MSFT ego network increased proving that investors focused on smaller group of well established stocks, hence creating stronger connections between them. Post the pandemic the average degree of MSFT remained relatively high proving that even though investors were expanding their portfolios they were still being cautious.

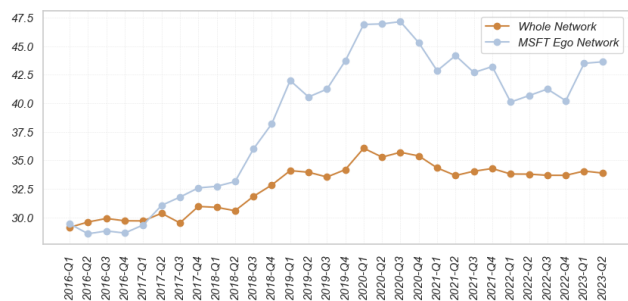


Figure 14: Average Strength in the Bipartite Ego Network of MSFT from 2016 to 2023.

The average strength indicates the average weight of the connections, which shows how often investments are made over time in MSFT by its connected investors. At the beginning the average strength of both the networks was similar showing that investors made consistent investments across stocks. During the pandemic, the average strength of the MSFT network increased, proving that investors believed in MSFT related stocks and invested more in it as it made them feel safe investing in familiar stocks in uncertain times. After the pandemic the average strength remained high proving that the investors were confident in MSFT related stocks.

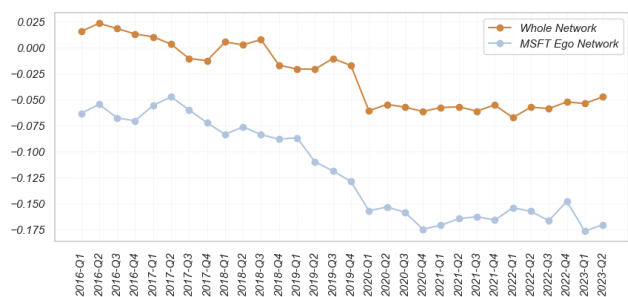


Figure 15: Assortativity in the Bipartite Ego Network of MSFT from 2016 to 2023.

Assortativity measures the tendency of nodes to connect to others with similar properties, providing insights into how homogeneous or heterogeneous the network is over time. Both networks showed slightly negative assortativity, proving well-connected stocks were more likely to be linked to lesser known stocks. During the pandemic the assortativity of MSFT network fell even more which indicated that investors would often complement a popular stock with a lesser known stock to balance its risk. Post the pandemic it increased very slightly trying to balance the investors portfolios as they included popular as well as new emerging stocks

Task 4.2 - Centrality

The influence of different stocks in the network was studied by using a centrality measure that highlights which stocks are strategically important. We chose betweenness centrality because it shows which nodes act as bridges in the network, essentially

connecting different parts of the market. It tells us which stocks are the key players in linking diverse investor interests—making them particularly influential in the market. Betweenness centrality is especially insightful for financial analysis since it points to where investor sentiment and focus are most concentrated. Stocks that rank high in betweenness are likely to control the flow of interest and information, making them central to market dynamics. Every quarter, we used betweenness centrality to identify the top three most central stocks and compared them to Microsoft’s (MSFT). Table 5, highlights MSFT’s centrality scores compare with the top three stocks for each quarter from 2016 to 2023. Unsurprisingly, big names like Google (both GOOG and GOOGL), Meta (formerly Facebook), and others like Berkshire Hathaway and Wells Fargo frequently showed up among the top ranks. It is interest to note that MSFT wasn’t always the most central stock even though it was always at the top of the market. There were several times—like in the third quarter of 2017 or the fourth quarter of 2020—where Google had a much higher score, reaching up to 0.114, while MSFT’s scores stayed between 0.044 and 0.056. This tells us that although MSFT has always been a significant player, it wasn’t always the main bridge that connected investors to other opportunities in the market. Figure 16 helps visualise these patterns by plotting the centralities over time. The solid peru coloured line represents MSFT, while the dashed lines show the top three most central stocks each quarter. MSFT’s centrality is quite stable, generally ranging between 0.04 and 0.08, while the other top stocks have much more fluctuation, sometimes spiking above 0.10. During the pandemic, MSFT’s centrality held steady, which suggests that investors saw it as a safe, reliable choice. Meanwhile, stocks like Google had big spikes in their centrality, like when it hit 0.107 in Q4 2020—maybe reflecting a moment when investors were looking for opportunities in a turbulent market.

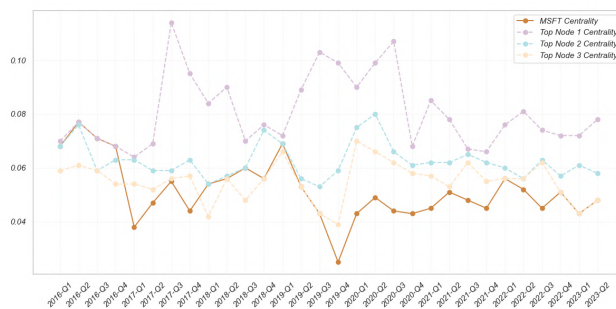


Figure 16: Comparison of MSFT Centrality Over Time with the Top 3 Most Central Nodes

Although MSFT has consistently been a key stock, it wasn’t always the most influential. Other companies, particularly Google and Meta, took on more central roles at different points, which likely reflected their appeal during specific market events or trends. MSFT’s steadiness, on the other hand, implies it was viewed as a dependable investment—something that stayed relevant no matter what else was happening in the market. For DBBA Capital, this information could be really useful. MSFT

Year	Quarter	Top 3 Nodes	Top 3 Centralities	MSFT Centrality
2016	1	[GOOG, MSFT, BRK.B]	[0.07, 0.068, 0.059]	0.068
2016	2	[MSFT, GOOG, AGN]	[0.077, 0.076, 0.061]	0.077
2016	3	[MSFT, GOOG, AAPL]	[0.071, 0.059, 0.059]	0.071
2016	4	[MSFT, BRK.B, GOOG]	[0.068, 0.063, 0.054]	0.068
2017	1	[BAC, GOOG, WFC]	[0.064, 0.063, 0.054]	0.038
2017	2	[GOOG, BAC, AABA]	[0.069, 0.059, 0.052]	0.047
2017	3	[GOOG, WFC, BAC]	[0.114, 0.059, 0.056]	0.055
2017	4	[GOOG, WFC, AAPL]	[0.095, 0.063, 0.057]	0.044
2018	1	[GOOG, MSFT, LEN]	[0.084, 0.054, 0.042]	0.054
2018	2	[GOOG, C, MSFT]	[0.09, 0.057, 0.056]	0.056
2018	3	[GOOG, MSFT, BRK.B]	[0.07, 0.06, 0.048]	0.060
2018	4	[META, GOOG, MSFT]	[0.076, 0.074, 0.056]	0.056
2019	1	[META, MSFT, GOOG]	[0.072, 0.069, 0.066]	0.069
2019	2	[GOOG, META, MSFT]	[0.089, 0.056, 0.053]	0.053
2019	3	[GOOG, META, MSFT]	[0.103, 0.053, 0.043]	0.043
2019	4	[GOOG, META, GOOGL]	[0.099, 0.059, 0.039]	0.025
2020	1	[META, GOOG, GOOGL]	[0.09, 0.075, 0.070]	0.043
2020	2	[GOOG, META, GOOGL]	[0.099, 0.080, 0.066]	0.049
2020	3	[META, GOOGL, GOOG]	[0.107, 0.066, 0.062]	0.044
2020	4	[META, WFC, GOOGL]	[0.068, 0.061, 0.058]	0.043
2021	1	[META, GOOGL, GOOG]	[0.085, 0.062, 0.057]	0.045
2021	2	[META, GOOG, GOOGL]	[0.078, 0.062, 0.053]	0.051
2021	3	[GOOG, META, GOOGL]	[0.067, 0.065, 0.062]	0.048
2021	4	[GOOGL, GOOG, META]	[0.066, 0.062, 0.055]	0.045
2022	1	[META, GOOG, MSFT]	[0.076, 0.060, 0.056]	0.056
2022	2	[GOOG, GOOGL, META]	[0.081, 0.056, 0.056]	0.052
2022	3	[GOOG, GOOGL, META]	[0.074, 0.063, 0.062]	0.045
2022	4	[GOOG, GOOGL, MSFT]	[0.072, 0.057, 0.051]	0.051
2023	1	[GOOG, GOOGL, MSFT]	[0.072, 0.061, 0.043]	0.043
2023	2	[GOOG, GOOGL, MSFT]	[0.078, 0.058, 0.048]	0.048

Table 5: Top 3 Nodes, Centralities, and MSFT Centrality by Year and Quarter

seems to provide stable, consistent value, which makes it a good anchor for an investment portfolio, especially during uncertain times. However, the fact that other stocks like Google or Meta occasionally spiked in influence suggests there might be opportunities to capitalize on those shifts. Recognizing which stocks are becoming more central at different times could help DBBA Capital make more strategic decisions, balancing the reliability of MSFT with the dynamic opportunities presented by these other key players. For DBBA Capital, understanding these shifts is key—it could mean knowing when to lean on MSFT for stability and when to ride the wave with other stocks that are gaining influence.

Task 5 - Clustering and Modularity

Louvain community detection algorithm was used to identify communities in the stock-side network projection for each quarter. This method was chosen as it is efficient in handling large datasets as well as graphs and it focuses on modularity optimisation. Since the dataset spans over many years and has

large number of stocks, using this algorithm was important to identify communities with high modularity. Clusters of stocks that have many international connections but have few links to others are identified. This helps understand how different stocks, like MSFT have grouped together and evolved over time. The main aim is to highlight MSFT’s community changes across the quarters and whether MSFT was consistently part of the same community with similar stocks or if there were noticeable changes. The MSFT’s community had variations across quarters the quarters and this could be due to various market circumstances, investor sentiment and other economic conditions. Before COVID, between 2016 and 2019, Microsoft’s stock community was considerably stable and predictable. It included well known companies, especially in tech and finance. This suggested how investors were thinking—they wanted steady, high-growth options and focused on the more popular stocks. Hence a lot of the investments of Microsoft paired up with companies like Google, Amazon, and JPMorgan, because these were the trusted stocks for anyone wanting stability and a solid return. During the pandemic (March 2020 to March 2021), MSFT’s community of connected stocks grew significantly—including

as many as 168 stocks, which was much larger compared to the pre-pandemic average of around 102. This jump indicated that investors were looking for security, and Microsoft became one of their "safe harbors" in these uncertain times. Although investors diversified their portfolios, they still anchored them to Microsoft, to lower their risk to reward ratio. This was a shift away from a more focused approach to one that prioritized security and stability, bundling a broad array of stocks alongside MSFT as investors sought a refuge in uncertain times.

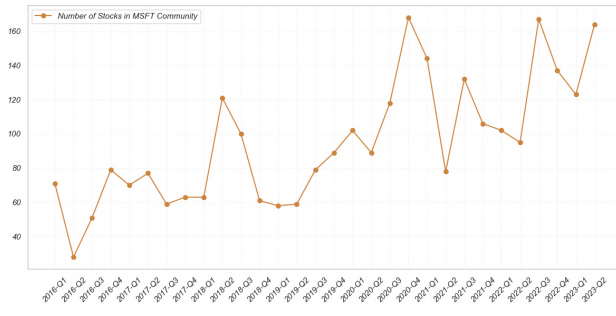


Figure 17: Number of Stocks in MSFT Community

After the pandemic, post March 2021, there was another change in the dynamics. The number of stocks connected to Microsoft began to fluctuate more. While the community was still diverse, it wasn't as broad as during the peak of the pandemic. Other big players like Google and Meta often showed up alongside Microsoft, but investors were starting to broaden their portfolios. Modularity also increased during this time, which essentially means investors were focusing back on specific sectors rather than spreading out too broadly as they did in peak crisis time, and hence focused on targeted investing. Throughout all these changes, Microsoft remained grouped with key companies like Google, Meta, and Amazon. This pattern showed that these companies were central in connecting a wide array of investor interests. They were like the glue holding together different pieces of the market. But, interestingly, the number and variety of stocks which were a part of Microsoft's community varied depending on the economic climate. During times of crisis, the community expanded, showing that investors leaned heavily on Microsoft, Google and Meta for stability. Conversely, as stability returned, the community became more selective, with investors concentrating their focus.

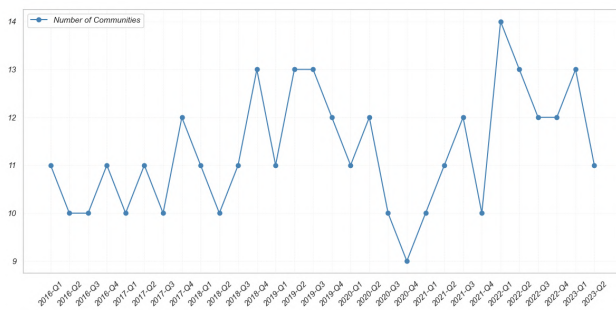


Figure 18: Number of Communities over quarters

In Figure 18 the graph portrays the number of communities across quarters. It can be noted that the number of communities have stayed pretty consistent ranging from 10 to 13 with a few peaks and drops. At the beginning of 2021 during the pandemic the number of communities dropped significantly, this brought stocks closer together. However at the end of 2021 the number of communities peaked highlighting more prominent clusters of stocks within the network.

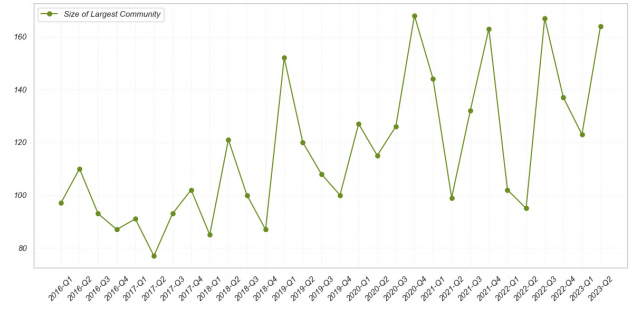


Figure 19: Size of the Largest Communities over quarters

The size of the largest community across quarters in portrayed in Figure 19. This provides insight into the modularity of the network overtime. The size of the largest community grew during the pandemic, around early 2021 when it peaked to 160. This indicates that investors that investors started strengthening their portfolios in times of uncertainty, which in turn strengthened the bond of certain clusters. After the pandemic investors once again started branching out into newer stocks to diversify their portfolios and reduced the size of the largest communities.

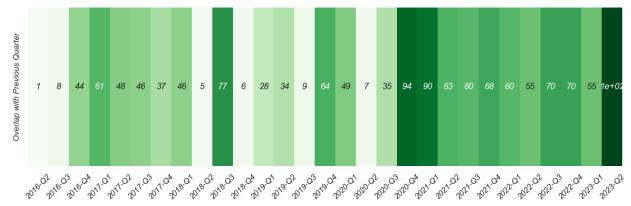


Figure 20: Heatmap of the community overlaps across quarters

Figure 20 highlights the stability of the communities across quarters. Darker green means there was a lot of overlap, indicating stability, while lighter shades show a shift in the community compositions. In the early years of 2016-2017 we can notice minimal overlap with past quarters indication frequent changes in the community. In 2017 Q1 as well as 2018 Q3 there was a noticeable overlap proving that the community had remained consistent and stable for that period with investors holding onto similar stocks for consecutive quarters. There is a high overlap in 2021-Q4 and 2023-Q2 as well to emphasize the stability. Table 6 complements these findings providing more details and exact figures. It shows how during the pandemic, MSFT's community grew significantly and after the pandemic, MSFT was still grouped with major players like Google and Meta, but the community became more selective, with fewer overall connections. This shift suggests that investors moved from broad

Period	MSFT Community	Stocks in MSFT Community	Number of Communities	Size of Largest Community	Number of Stocks in MSFT Community
2016 Q1	3	GE, WU, ZBH, ZBRA, MSFT, ...	11	97	71
2016 Q2	3	MSFT, TTE, CACI, BHI, EOG, ...	10	110	28
2016 Q3	0	COF, MSFT, STLA, ORCL, CBRE, ...	10	93	51
2016 Q4	3	GOOG, KKR, JPM, BAC, C, ...	11	87	79
2017 Q1	4	GOOG, JPM, BAC, C, MSFT, ...	10	91	70
2017 Q2	2	JPM, C, LVL, MSFT, STLA, ...	11	77	77
2017 Q3	4	GOOG, JPM, BAC, STLA, C, ...	10	93	59
2017 Q4	4	JPM, BFH, STLA, BAC, C, ...	12	102	63
2018 Q1	3	CBRE, JPM, BFH, C, PG, ...	11	85	63
2018 Q2	1	BRK.B, AMZN, GOOG, CVNA, BAX, ...	10	121	121
2018 Q3	3	BRK.B, AMZN, GOOG, CVNA, BAX, ...	11	100	100
2018 Q4	11	BDX, MSFT, MDLZ, ECL, PG, ...	13	87	61
2019 Q1	4	ORCL, MSFT, BRK.A, PG, ELV, ...	11	152	58
2019 Q2	3	MSFT, AAPL, PG, ADPT, ORCL, ...	13	120	59
2019 Q3	8	MSFT, CVNA, BAX, SE, UNP, ...	13	108	79
2019 Q4	5	JD, CVNA, BAX, MSFT, SE, ...	12	100	89
2020 Q1	1	OI, AMZN, GOOGL, JD, CVNA, ...	11	127	102
2020 Q2	1	LIN, AMT, DIS, MSFT, MDLZ, ...	12	115	89
2020 Q3	0	LIN, CHTR, GOOGL, CVNA, MA, ...	10	126	118
2020 Q4	0	LIN, TDG, CBRE, CHTR, MA, ...	9	168	168
2021 Q1	0	LIN, GOOGL, DXC, CHTR, JD, ...	10	144	144
2021 Q2	7	JD, CVNA, UPST, MSFT, SE, ...	11	99	78
2021 Q3	2	CHTR, AMZN, MSFT, GOOG, MA, ...	12	132	132
2021 Q4	2	CHTR, AMZN, JD, GOOG, CVNA, ...	10	163	106
2022 Q1	4	AMZN, JD, GOOG, MA, CVNA, ...	14	102	102
2022 Q2	10	GOOGL, JD, GOOG, MSFT, CP, ...	13	95	95
2022 Q3	0	BKNG, TDG, MSFT, GOOGL, JD, ...	12	167	167
2022 Q4	0	BKNG, TDG, MSFT, BAC, BRK.B, ...	12	137	137
2023 Q1	3	BKNG, TDG, MSFT, GOOGL, META, ...	13	123	123
2023 Q2	0	BKNG, TDG, MSFT, GOOGL, META, ...	11	164	164

Table 6: Community analysis of MSFT and associated stocks over time (top 5 stocks shown).

diversification back to a more targeted focus. These insights can prove to be very strategic for DBBA capital for the following reasons: During period of high market uncertainty, the expansion of Microsoft’s community can serve as a signal to diversify more broadly and strengthen positions in stable companies like MSFT which creates a safety net and reduces risk. On the other hand, when the market is more stable and MSFT’s community shrinks, it could be beneficial to focus more on emerging trends or sector-specific plays thereby capturing high potential growth opportunities. Thus, by carefully observing such trends, DBBA Capital can adapt its strategies to match the market mood. This also helps them to streamline their decisions and understand when it makes sense to diversify stocks and protect against risk, and when it’s better to narrow in on specific sections to capitalise on high returns. The way in which they align their investment strategies to the evolving dynamics of the market can allow DBBA Capital to stay ahead of market trends and optimise their position in various situations- either by casting a wide net or honing in on a particular segment.

Task 6 - Analysing Stock Correlation Through a Spanning Tree

Task 6.1 - Construct a Stock Correlation Network

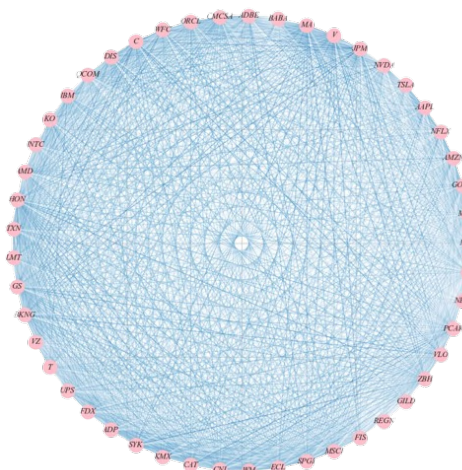


Figure 21: Stock Correlation matrix.

Using the provided closing stock prices dataset a stock correlation network was created, where each stock is a node, and edges are weighted based on the correlations between the monthly returns of each pair of stocks. The network consists of 49 nodes and 1176 edges, proving the network to be extremely interconnected. The visualisation in Figure 18 portrays a complex structure, where the pink node represents the stocks and the edges show how interlinked these stocks are. This helps observe trends and patterns in the market.

Task 6.2 - Build a Spanning Tree

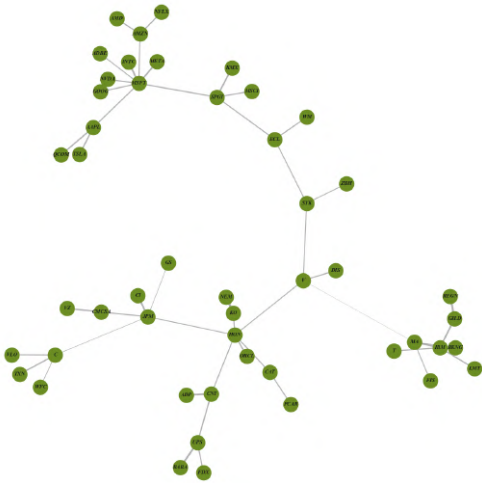


Figure 22: *Minimum Spanning Tree of Stock Correlation Network*

Top 5 edges by distance			Bottom 5 edges by distance		
Stock 1	Stock 2	Distance	Stock 1	Stock 2	Distance
V	MA	0.080957	KO	NEM	0.714691
JPM	C	0.164608	REGN	GILD	0.601491
JPM	GS	0.177837	BABA	UPS	0.549906
WFC	C	0.218869	MSFT	INTC	0.531895
SYK	ZBH	0.267813	JPM	CI	0.522365

Table 7: *Top and Bottom 5 Edges by Distance*

Using the correlation matrix constructed in 6.1 a minimum spanning tree algorithm was applied to it, represent the relationship among the stocks. This algorithm reduces complexity within the networks and displays only essential links between the stocks. In this plot each node represents a stock and the edge between them is weighted by distance. To calculate this distance the formula is $1 - \text{abs}(\text{correlation})$, which indicates a lower distances correspond to stronger connections. This method highlights the strongest connections while filtering out the weaker, noisier ones, making it easier to see meaningful patterns. In the plot different clusters are identified which show they have a close relationship among each other. The thickness of the edge

indicates the strength of these relations. Strong correlations can be spotted. Tables 7 and 8 provide a summary of the top and bottom five connections based on distance. Stocks like KO and NEM do not tend to move together and show weaker connections due to the distance between them. Market behavior is very similar in stocks like V and MA that have a smaller distance between them and are closely linked.

Task 6.3 - Analyse MSFT's Role in the Spanning Tree

MSFT's position within the spanning tree was analyzed by discussing its importance in terms of centrality and its connections to other stocks before, during, and after the pandemic. Microsoft had a moderately connected network and was linked with big companies such as Adobe Amazon and Canadian National Railway before March 2020. Its degree centrality of 0.0625 showed that it had connections, but wasn't necessarily dominating the scene. Microsoft had a decent role and was reliable in bridging different clusters of stocks, it has a betweenness centrality of 0.1587. Its closeness centrality of 0.1951 meant it wasn't far from the other stocks in the network, making it pretty well-positioned. This highlights tha Microsoft was a trusted and well connected stock but was not the center most stock at the beginning of the pandemic.

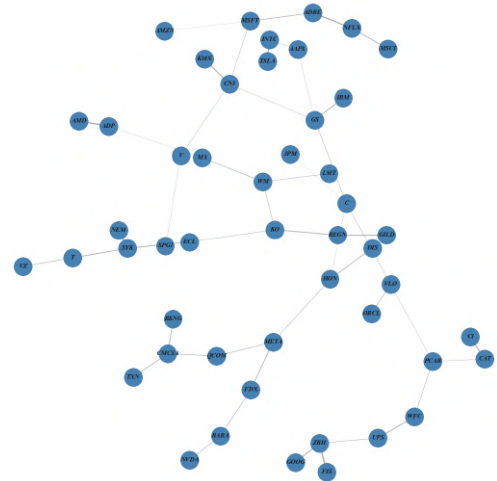


Figure 23: *Minimum Spanning Tree of Stock Correlation Network Before the Pandemic*

The market shifted dramatically once the pandemic began (March 2020 to March 2021) which impacted Microsoft's influence in the network. Its degree centrality dropped to 0.0283, and it was only directly connected to Tesla. This huge drop in connections hints that during the pandemic, when uncertainty was high, investors didn't lean as heavily on Microsoft. Most notably, betweenness centrality fell all the way to 0, meaning Microsoft was no longer playing that crucial "connector" role it previously had. Investors were likely scrambling, spreading out their portfolios, and Microsoft's role as a bridge between various parts of the market just disappeared. Even closeness

Period	Degree Centrality	Betweenness Centrality	Closeness Centrality	Neighbours in MST
Before Pandemic	0.062500	0.158688	0.195122	[AMZN, CNI, ADBE]
During Pandemic	0.020833	0.000000	0.136752	[TSLA]
After Pandemic	0.083333	0.577128	0.211454	[GOOG, SPGI, MSCI, TSLA]

Table 8: MSFT Centrality Measures and Neighbours in the Spanning Tree Before, During, and After the Pandemic

centrality dropped to 0.1368, implying Microsoft was further from the rest of the stocks in the network. It seems that during the pandemic, people were more cautious, diversifying in a way that left Microsoft out of the central picture.

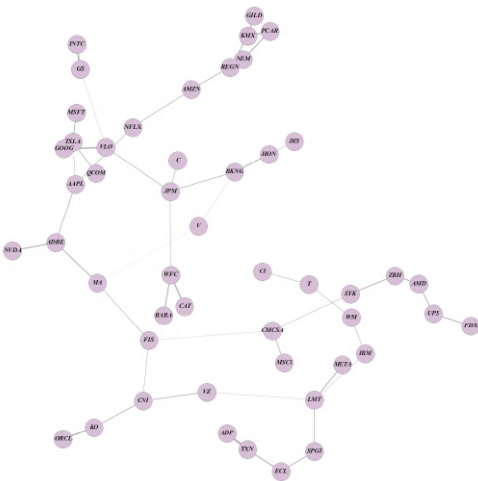


Figure 24: Minimum Spanning Tree of Stock Correlation Network During the Pandemic

After March 2021, Microsoft made a comeback. Its degree centrality went up to 0.0833, which, while still not massive, showed it was making more direct connections with other stocks, like Google, S&P Global, MSCI, and Tesla. The big change was in betweenness centrality, which jumped to 0.5771. Microsoft was back as an important bridge, connecting different parts of the market. This tells us that investors were once again viewing Microsoft as a stable anchor for their strategies. Closeness centrality also increased to 0.2115, meaning it was closer to the rest of the stocks again. Essentially, as the market started recovering, Microsoft regained its central role—investors were trusting it as a reliable point of connection. Overall, Microsoft’s journey through these periods highlights investor behavior and market sentiment. Before the pandemic, Microsoft was reliable but not pivotal. The plot highlights how Microsoft’s role shifted throughout the different stages of the pandemic. Before the pandemic, Microsoft had moderate influence, but during the pandemic, betweenness centrality (the purple line) drop almost to zero. This means Microsoft wasn’t really playing a big role in connecting other companies during the chaotic times. But after the pandemic, it bounces back up sharply, showing that Microsoft once again became a crucial bridge between different parts of the network. Degree centrality (the blue line) stayed pretty low throughout, meaning Microsoft didn’t necessarily

have a lot of direct connections compared to others. And closeness centrality (the orange line) remained fairly stable, with a slight increase after the pandemic, suggesting that Microsoft became easier to reach in the network as things settled down. This graph really captures how Microsoft lost some of its network influence during the pandemic but quickly regained it as stability returned.

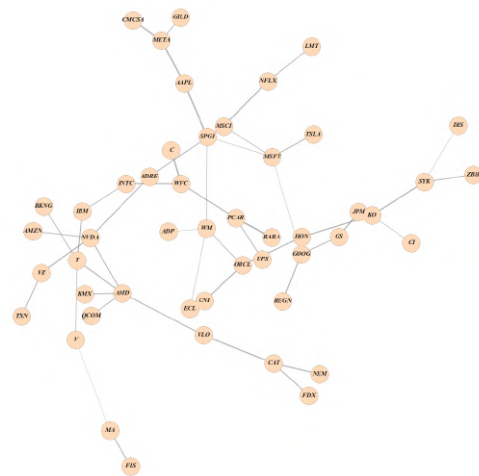


Figure 25: Minimum Spanning Tree of Stock Correlation Network After the Pandemic

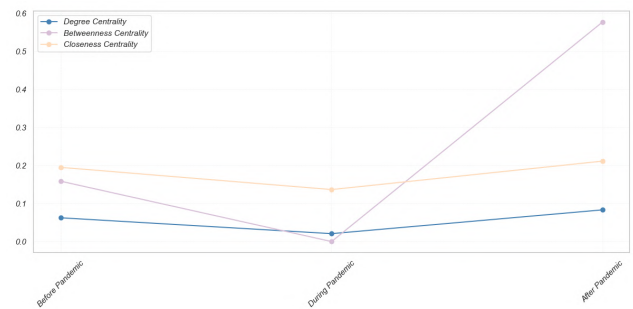


Figure 26: MSFT Centrality Analysis within the Spanning Tree

Task 7 - Result Discussion

In the above 6 task, MSFT stock has been analysed and it makes a strong case for DBBA Capital to invest in it. MSFT’s performance before, during, and after the pandemic makes it a stable and reliable stock. Microsoft consistently remained a trusted anchor for investors, even during market upheaval. Before the

pandemic, MSFT was well-connected, though not always the top influencer. During the pandemic, its role as a market bridge diminished, reflecting investors' shift to diversify. However, post-pandemic, MSFT rebounded strongly, regaining its role as a key connector in the market. Community analysis showed that during the pandemic, MSFT's community expanded, indicating it to be a safe stock for many investors. Post-pandemic, while the community became smaller, MSFT remained influential, especially among tech stocks. This demonstrates the trust investors have in its long-term value. The spanning tree analysis further highlights MSFT's adaptability. Its connections dropped during the pandemic but bounced back afterward, showing its ability to recover and stay relevant. To sum it up, I recommend DBBA Capital invest in Microsoft (MSFT). The evidence from our analyses—network connections, spanning tree metrics, and community changes—shows that MSFT is a dependable, influential player in the market. It consistently provides value, whether the market is facing challenges or heading into growth mode. Its combination of being a safe bet and a connector in the broader market makes it an ideal choice for a well-rounded, strategic investment.